

Hardness and Approximation for Network Flow Interdiction

Stephen R. Chestnut* Rico Zenklusen†

November 10, 2015

Abstract

In the Network Flow Interdiction problem an adversary attacks a network in order to minimize the maximum s - t -flow. Very little is known about the approximability of this problem despite decades of interest in it. We present the first approximation hardness, showing that Network Flow Interdiction and several of its variants cannot be much easier to approximate than Densest k -Subgraph. In particular, any $n^{o(1)}$ -approximation algorithm for Network Flow Interdiction would imply an $n^{o(1)}$ -approximation algorithm for Densest k -Subgraph. We complement this hardness results with the first approximation algorithm for Network Flow Interdiction, which has approximation ratio $2(n-1)$. We also show that Network Flow Interdiction is essentially the same as the Budgeted Minimum s - t -Cut problem, and transferring our results gives the first approximation hardness and algorithm for that problem, as well.

Keywords: Network flow interdiction, approximation algorithms, hardness of approximation, budgeted optimization

1 Introduction

We are given an undirected graph $G = (V, E)$ with edge capacities $u(e) \geq 0$, for all $e \in E$, and distinct vertices $s, t \in V$. An adversary removes edges from the graph with the goal of reducing the maximum s - t -flow. It costs the adversary $c(e)$ to remove edge e and he has a total budget B for removing edges. The *Network Flow Interdiction (NFI)* problem is to determine the optimal strategy for the adversary. More precisely, given G , s , t , u , c , and B , the goal is to find a set of edges $R \subseteq E$ such that $c(R) := \sum_{e \in R} c(e) \leq B$ and the maximum s - t -flow in $(V, E \setminus R)$ is minimized.

Network flow interdiction has a long history in combinatorial optimization beginning with the famous max-flow/min-cut theorem of Ford and Fulkerson. A declassified RAND report from 1955 [13] regarding interdiction of the Soviet rail network in Eastern Europe is cited by Ford and Fulkerson as motivation for the minimum s - t -cut problem [22]. If there is an s - t -cut with cost at most B , then an optimal strategy for the adversary is to remove all of its edges. On the other hand, if all of the s - t -cuts have cost greater than B then there need be no relation between min cost cuts and optimal NFI solutions. Besides the

*Department of Mathematics, ETH Zurich. E-mail: stephenc@ethz.ch.

†Department of Mathematics, ETH Zurich, and Department of Applied Mathematics and Statistics, Johns Hopkins University. E-mail: ricoz@math.ethz.ch.

obvious military applications like supply-line disruption, Network Flow Interdiction and its variants have been proposed as a model for combating the spread of infectious diseases in a hospital [1], drug interdiction [23], and critical infrastructure analysis [18]. NFI assesses the worst case impact of a limited set of failures in a network flow, so it can also be used to assess the robustness of a network.

Despite decades of interest in the problem, surprisingly little is known about its approximability. Only strong NP-hardness is known by reductions from Clique and Minimum Bisection [23, 20], and also the Budgeted Minimum s - t -Cut (BMstC) problem, which we show is essentially equivalent to the NFI problem, has been proved to be strongly NP-hard [19].

Several authors have proposed exact mixed-integer programming formulations and enumeration based algorithms for NFI and its variants (see [23] and [21] and the references therein). But as far as polynomial time algorithms go, there is only a $(1 + \epsilon, 1 + \frac{1}{\epsilon})$ -pseudoapproximation algorithm due to Burch et al. [7]. Specifically, for any $\epsilon > 0$, that algorithm returns either a $(1 + \frac{1}{\epsilon})$ -approximation, or a super-optimal solution that violates the budget by as much as a factor of $1 + \epsilon$; however, one cannot choose which of the two outcomes happens. In the special case when G is a planar network, there is also an FPTAS [20], which can be extended to allow for removal of vertices [25]. We begin to fill this gap on the hardness side by showing that NFI cannot be much easier to approximate than the Densest k -Subgraph (DkS) problem, and on the approximation side by giving the first true approximation algorithm for NFI, which has approximation ratio $2(n - 1)$. We also supply an approximation preserving reduction from BMstC to NFI and a reverse reduction that preserves the approximation ratio up to a factor of $(1 + \epsilon)$. This allows us to use our approximation algorithm for NFI as a $2(n - 1)$ -approximation algorithm for BMstC and proves that BMstC cannot be much easier to approximate than DkS.

The DkS problem is to determine for a given graph H the maximum number of edges in any k -vertex subgraph of H . DkS is clearly NP-hard since it captures the problem of finding a clique of size k . The complexity of approximating DkS is still an open problem, but the available evidence indicates that the problem could be very hard to approximate. The first polynomial time approximation algorithm, given by Feige, Peleg, and Kortsarz in 1997 [11], has approximation ratio $O(n^{1/3})$, and this was not improved until 2010 when Bhaskara, Charikar, Chlamtac, Feige, and Vijayaraghavan [5] found an algorithm with approximation ratio $O(n^{\frac{1}{4} + \epsilon})$, for any $\epsilon > 0$. The latter is the currently best known guarantee. As for hardness, under different complexity-theoretic assumptions Feige [10] (assuming random 3-SAT instances are hard to refute) and Khot [16] (assuming NP does not have sub-exponential time randomized algorithms) have shown that there is no PTAS for DkS. This extends to our results in that the same assumptions imply that no PTAS exists for NFI. Returning to DkS, significant gains have been ruled out for lift and project hierarchies. An integrality gap of $\Omega(n^\epsilon)$ persists even for $n^{1-O(\epsilon)}$ rounds in the Lasserre hierarchy, which implies the same for Sherali-Adams and Lovász-Schrijver hierarchies, and for the Sherali-Adams hierarchy the integrality gap is $\Omega(n^{1/4}/\log^3 n)$ after $O(\frac{\log n}{\log \log n})$ rounds [6], matching the best known approximation ratio.

It is often easy to prove NP-hardness for interdiction variants of combinatorial optimization problems, but our result is one of only a handful of approximation hardness theorems for interdiction problems. Others include APX hardness of shortest path interdiction by removal of edges [15], k -median and k -center interdiction [2], and assignment problem interdiction [3] as well as clique hardness for spanning tree interdiction by removal of nodes [4]. Approximation or bicriteria approximation algorithms are known for the interdiction ver-

sions of minimum spanning tree by removal of edges [12, 26], (fractional) multicommodity flow [8], matching [24] and more generally packing LPs [9]. Overall, strikingly little is known about interdiction variants of common optimization problems, despite many natural applications.

The next section presents our notation and precise definitions of the problems. Approximation hardness for NFI is proved in Section 3, and Section 4 describes the $2(n - 1)$ -approximation algorithm. Section 5 establishes connections between NFI and several of its variants, as well as with BMstC.

2 Preliminaries

Let $G = (V, E)$ be an undirected graph which may have parallel edges. We often use G to represent an NFI instance. A set of vertices $C \subseteq V$ is an s - t -cut if $s \in C$ and $t \notin C$ and we sometimes also identify a cut by its set of edges $\delta_G(C) = \{e \in E \mid e \text{ has exactly one endpoint in } C\}$. When the graph is clear from the context we may drop the G subscript. The cost of a cut is $c(\delta(C))$ and its capacity is $u(\delta(C))$. H is also an undirected graph that usually represents the input for DkS. Given $S \subseteq V(H)$ the density of S is $d_H(S) = |E_H[S]|/|S|$, where $E_H[S]$ is the set of edges of H with both endpoints in S . Similarly, we use $E[S, T]$, for $S \cap T = \emptyset$, to denote the set of edges with one endpoint in S and one in T .

Let $\mathbb{N} = \{0, 1, 2, \dots\}$. An instance of NFI is an undirected graph $G = (V, E)$, capacities $u : E \rightarrow \mathbb{N}$, distinct vertices $s, t \in V$, costs $c : E \rightarrow \mathbb{N}$, and a budget B , and the goal is to minimize the value of a maximum s - t -flow in $(V, E \setminus R)$ by choosing a set $R \subseteq E$ satisfying $c(R) \leq B$. An instance of BMstC is the same as for NFI and the goal is to determine the minimum capacity $u(C)$ of any s - t -cut $C \subseteq V$ such that $c(\delta(C)) \leq B$. An instance of DkS is a simple, undirected graph H and an integer $0 < k < n$ and the goal is to identify a set of k vertices $W \subseteq V(H)$ that maximizes the number of edges (equivalently the density) in the subgraph of H induced by the vertices W , which we denote by $H[W]$.

An algorithm is an α -approximation for a given minimization problem if it is a polynomial time algorithm that produces a feasible solution with objective no more than $\alpha \cdot \text{OPT}$. For a maximization problem, the objective value is at least OPT/α .

For convenience, we often use ∞ as the cost or capacity of an edge. This is merely a notational contrivance, it does not affect the generality since the cost of any edge can be decreased to $B + 1$ without affecting the problem, and if an instance with infinite capacity edges has finite maximum flow then the capacities can be reduced to this value. Setting the cost of an edge to infinity is just a way to denote an edge that can never be removed. Similarly, an infinite capacity edge can never be in a minimum s - t -cut after interdiction.

3 Hardness of approximation

The form of our reduction from DkS to NFI is inspired by two other hardness proofs, one published in [23], which proves NP-hardness of NFI on directed graphs, and another proof from [25], which reduces DkS on planar graphs to NFI with vertex deletions on directed planar graphs with multiple sources and sinks, although neither considers hardness of approximation. In both works, the authors subdivide the edges of the DkS input graph H , direct the new edges, and use a series of NFI instances with different budgets in order to solve either Clique or DkS, respectively. The main challenge in proving DkS-hardness for

NFI is to design new edge costs and capacities that lead to well-structured interdiction solutions.

Let us first outline the strategy of the reduction. Given a DkS instance $H = (V, E)$, we form a new graph G with $|V| + |E| + 2$ vertices and $|V| + 3|E|$ edges by subdividing the edges of H and adding a source and a sink, as shown in Figure 1. With the right choice of B , s - t -flow interdiction solutions in G correspond to subgraphs of H that have the same number of edges as a densest k -subgraph. Furthermore, the optimal interdiction solution corresponds exactly to a densest k -subgraph, and an α -approximate interdiction solution yields a subgraph with at most αk vertices. Such a subgraph K with size between k and αk easily certifies the existence of a k -subgraph with density at least $\frac{1}{2\alpha^2}d(K)$ in the original graph; it suffices to consider a random induced subgraph of K with k vertices. Furthermore, one can also easily find such a high-density subgraph of K deterministically, as we will discuss briefly later.

The upshot is a factor $2\alpha^2$ approximation algorithm for DkS where α is the NFI approximation ratio for an input graph with $|V| + |E| + 2 \leq |V|^2$ vertices. The algorithm returns the approximate densest k -subgraph as long as the interdiction algorithm returns the set $R \subseteq E$ of interdicted edges, otherwise it just approximates the objective value.

The next lemma establishes the second step in our reduction, namely given a graph with ℓ^* edges, where ℓ^* is the number of edges in a densest k -subgraph of H , and at most αk vertices it finds a k -subgraph with at least $\ell^*/2\alpha$ edges. A similar lemma appears in [14], but the algorithm we present is simpler and more efficient and we include it for completeness. The proof that NFI is roughly as hard to approximate as DkS follows it.

Lemma 1. *There is a deterministic algorithm that takes as input a graph $H = (V, E)$, with $|V| \geq k$, and returns an induced subgraph K of H with k vertices and density at least $\frac{k-1}{|V|-1}d(H)$. The running time of the algorithm is $O(|E|)$.*

Proof. We begin with a randomized algorithm that achieves density at least $d = d(H) \frac{k-1}{|V|-1}$ in expectation and proceed to derandomize it with the method of conditional expectations. The randomized algorithm chooses a uniformly random set of k vertices from $V(H)$ and takes K to be the induced subgraph on these vertices. Defined this way, the expected number of edges in K is $\frac{k(k-1)}{|V|(|V|-1)}|E|$, hence the expected density is $\mathbb{E}[d(K)] = d$.

Let us label the vertices v_1, \dots, v_n arbitrarily and denote by V_i the set $\{v_1, v_2, \dots, v_i\}$. The derandomized algorithm is the following:

- 1 Initialize $S = \emptyset$
- 2 For $i = 1, 2, \dots, n$
- 3 If $|E[S + v_i]| + \frac{k-|S|-i}{n-i}|E[v_i, V \setminus V_i]| \geq \frac{1}{|V|-i}|E[S, V \setminus V_i]| + \frac{2(k-|S|-1)}{(|V|-i)(|V|-i-1)}|E[V \setminus V_i]|$
- 4 $S \leftarrow S \cup \{v_i\}$
- 5 Otherwise leave S unchanged
- 6 Return $K = (S, E[S])$

The inequality in Line 3 amounts to checking $\mathbb{E}[d(K) \mid V_i \cap K = S \cup \{v_i\}] \geq \mathbb{E}[d(K) \mid V_i \cap K = S]$. The algorithm is correct because it maintains the invariant $\mathbb{E}[d(K) \mid S \subseteq V(K)] \geq d$. To compute each of the values needed for a single iteration, given the values from the previous iteration, it suffices to pass once over the list of neighbors of v . Thus the total time required is $O(|E|)$. \square

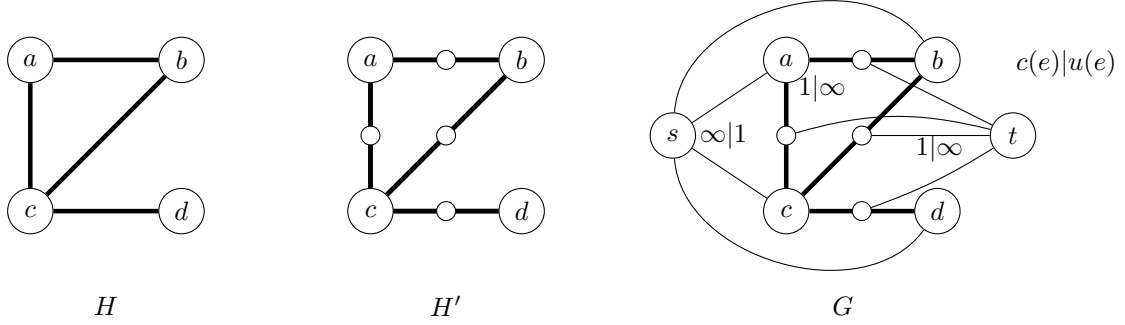


Figure 1: An example of the reduction from DkS to NFI. The DkS instance is given by the graph on the left and the graph on the right is the corresponding NFI instance with cost|capacity as follows $\delta(\{s\}) : \infty|1$, $\delta(\{t\}) : 1|\infty$, and the bold edges $1|\infty$.

Next, we will formally define the auxiliary graph with its cost and capacity structure. Let $H = (V, E)$ be a DkS instance and form $H' = (V', E')$ by subdividing every edge of H with a new vertex. We associate the vertices of H' with the set $V \cup E$, i.e., $v \in V$ and $e \in E$ are both vertices of $V' = V \cup E$. We treat its edges similarly, each edge of E' is a pair $ve \in V \times E$ where e is incident with v in H . Let G be the graph with vertices $V' \cup \{s, t\}$ and edges $E' \cup \{sv \mid v \in V\} \cup \{et \mid e \in E\}$. That is, add source and sink vertices s and t , adjoin s with each V -vertex, and adjoin t with each E -vertex. Figure 1 presents an example of the completed construction. In G we assign costs and capacities as follows

Edge set	$\delta(s)$	E'	$\delta(t)$
cost c	∞	1	1
capacity u	1	∞	∞

It turns out that we may assume the interdiction solutions for G take a particularly simple structure. We call a set $R \subseteq E' \cup \delta(t)$ a *cut solution* if there exists a cut $C \subseteq V$ in the graph H such that

$$R = \{ve \in E' \mid v \in C \text{ and } e \in \delta_H(C)\} \cup \{et \mid e \in E_H[C]\},$$

and we use the notation $R(C)$ to denote the cut solution associated with the cut C . An example cut solution is shown in Figure 2. The NFI objective function value of a cut solution is $\max\text{-s-t-flow}(G - R(C)) = |V \setminus C|$, which is the number of vertices in H that are not in C , and the cost of a cut solution is

$$c(R(C)) = |\delta_H(C)| + |E_H[C]| = |E| - |E_H[V \setminus C]|.$$

A first consequence is that if C^* defines an optimal cut solution for NFI on G , then $H[V \setminus C^*]$ has the fewest vertices of any subgraph with $|E[V \setminus C^*]|$ edges, that is $H[V \setminus C^*]$ is a densest $|E[V \setminus C^*]|$ -edge subgraph. The next lemma shows that from any interdiction set R one can find a cut solution that has the same objective value and cost no more than $c(R)$.

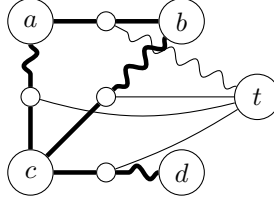


Figure 2: The cut solution defined by $C = \{a, b, d\} \subset V(H)$ in the auxiliary graph G of Figure 1. The four wavy edges form the cut solution.

Lemma 2. *There is a polynomial time algorithm that takes as input the graph G and a set of edges $R \subseteq E' \cup \delta(t)$ and returns a cut solution R' with $c(R') \leq c(R)$ and*

$$\text{max-s-t-flow}(G - R') \leq \text{max-s-t-flow}(G - R).$$

Proof. The algorithm applies successive transformations to R that preserve its objective value. Each transformation results in a net decrease in the cost of R . The main observation needed to evaluate the impact of the transformations is the following. Let $C \subseteq V'$ be the vertices of a connected component of $(V', E' \setminus R)$. Every maximum s - t -flow in $G - R$ assigns flow $|C \cap V|$ through C , unless $E[C, t] \subseteq R$, in which case the flow is zero. We call a component C a *flow component* or a *no-flow component* according as the flow through the C is $|C \cap V|$ or 0. The maximum flow in $G - R$ is $\sum_C |C \cap V|$, where the sum is taken over all of the flow components.

If there is an edge $ve \in R \cap E'$ whose endpoints lie in the same connected component of $H' - R$, then $R - ve$ is an interdiction solution defining the same flow and no-flow components, hence the same objective value. Similarly, given a flow component C we can remove from R any edge of $E[C, t]$, because, upon doing so, the flow through the component remains $|C \cap V|$ and the cost is decreased. The first step of the algorithm removes all of these edges from R .

Next, the algorithm “merges” any two components C_1 and C_2 of the same type by removing from R any edge in $E[C_1, C_2]$. This does not affect the objective value, because $ve \in V'$ is in a flow component before the deletion if and only if it is in a flow component after the deletion. Removing these edges from R is the second step of the algorithm.

Last, let C_0 be a no-flow component, let C_1 be a flow component, and let $ve \in R \cap E[C_0, C_1]$. Let $e = vw$ and two options appear. Either $v \in C_0$ and $w \in C_1$ or vice versa. No change is made if $v \in C_0$ and $w \in C_1$. If $v \in C_1$ and $w \in C_0$, then $e \in C_0$, which implies that $et \in R$ because C_0 is a no-flow component. In this case, the algorithm replaces R by $R - ve + we - et$. Doing so clearly reduces its cost. The new components of $G - R$ are $C_0 - e$, a no-flow component, and $C_1 + e$, a flow component. The flow through $C_1 + e$ is unchanged at $|(C_1 + e) \cap V| = |C_1 \cap V|$. Performing these changes is the last step of the algorithm. It returns as the set R' the modified set R . R' is the cut solution associated with the $C' = \cup_C (C \cap V)$, where the union is taken over all no-flow components of $G - R'$. \square

We are now in a position to complete the hardness proof. Let ℓ^* denote the number of edges in a densest k -subgraph of H . The reduction works by finding an approximate NFI interdiction solution with budget $|E| - \ell^*$, transforming it to a cut solution $R(C)$ with Lemma 2, interpreting $H[V \setminus C]$ an approximate densest ℓ^* -subgraph, and then sampling k -vertices from $V \setminus C$.

Theorem 3. *If there exists a polynomial time $\alpha(n)$ -approximation algorithm for NFI then there exists a polynomial time $2\alpha(n^2)^2$ -approximation algorithm for DkS, where in each case n is the number of vertices in the corresponding instance.*

Proof. Let $H = (V, E)$ be an instance of DkS. We first construct the auxiliary graph G with vertices $V' \cup \{s, t\}$ and edges $E' \cup \{sv \mid v \in V\} \cup \{et \mid e \in E\}$ as described above.

Here is the DkS approximation algorithm. For each $\ell = 1, 2, \dots, \binom{k}{2}$, run the NFI approximation algorithm on G with budget $B = |E| - \ell$ and let v_ℓ denote the residual flow value for this run. If $v_\ell \geq k$ then let $e_\ell = \ell \frac{k(k-1)}{v_\ell(v_\ell-1)}$ and otherwise let $e_\ell = \ell$. Return $\max_\ell e_\ell$ as the approximate number of edges in a densest k -subgraph.

Now we analyze the algorithm. Let $R_\ell \subseteq E' \cup \delta(t)$ be an interdiction set where $G - R_\ell$ has maximum s - t -flow equal to v_ℓ . By Lemma 2, we can assume that R_ℓ is a cut solution. Let $C_\ell \subseteq V$ be the associated cut, which has $|C_\ell| = v_\ell$ and $E[C_\ell] \geq \ell$, by the feasibility of R_ℓ . In the case of $v_\ell < k$, any k -subgraph containing C_ℓ has at least $E[C_\ell]$ edges. If $v_\ell \geq k$, then Lemma 1 implies that the induced subgraph $H[C_\ell]$ has itself a k -subgraph with density at least

$$\frac{k-1}{v_\ell-1} d(H[C_\ell]) \geq \frac{k-1}{v_\ell-1} \frac{\ell}{v_\ell} = \frac{e_\ell}{k}.$$

Therefore, for all ℓ , e_ℓ is a lower bound on the number of edges in some k -subgraph of H .

Let ℓ^* be the number of edges in a densest k -subgraph of H . The optimal objective value of the corresponding NFI problem when the budget is $|E| - \ell^*$ is at most k . Hence, $v_{\ell^*} \leq \alpha(|V|^2)k$ since G has $|V| + |E| + 2 \leq |V|^2$ vertices. Substituting into the definition of e_{ℓ^*} we have either $e_{\ell^*} = \ell^*$ or

$$e_{\ell^*} = \ell^* \frac{k(k-1)}{v_{\ell^*}(v_{\ell^*}-1)} \geq \frac{1}{2\alpha(|V|^2)^2} \ell^*, \quad (1)$$

according as v_{ℓ^*} is smaller or larger than k . This completes the proof of correctness. The running time bound follows immediately from the description of the algorithm. \square

Corollary 4. *Let $\epsilon > 0$. If there is a $\frac{1}{\sqrt{2}}n^{\epsilon/4}$ -approximation for NFI then there is a n^ϵ -approximation for DkS. In other words, any n^ϵ -approximation-hardness for DkS implies a $\frac{1}{\sqrt{2}}n^{\epsilon/4}$ -approximation-hardness for NFI.*

The factor 2 in Theorem 3 and $\sqrt{2}$ in the corollary can be reduced to any constant larger than 1 by assuming $k = \omega(1)$, which must hold for all hard instances, and adjusting the upper bound in (1).

If the NFI approximation algorithm in Theorem 3 also returns the interdiction sets $R_\ell \subseteq E' \cup \delta(t)$, then the algorithm from the last proof also produces an approximate densest k -subgraph. Indeed, by Lemma 2, we can find a cut solution from R_ℓ and the associated cut $C_\ell \subseteq V$. For each cut C_ℓ with at least k vertices we can use the algorithm of Lemma 1 to find a k -subgraph H with at least e_ℓ edges and return the largest subgraph among them.

Finally, it is worth pointing out that this reduction is efficient in its use of the edges in the graph. In deriving the approximation ratio, we have used the bound $m+n+2 \leq n^2$, but one also has $m+n+2 \leq 3m$. Hence, if there are hard instances of DkS with $m = o(n^2)$ edges then the gap between the two approximation ratios in Theorem 3 is narrowed to $\alpha(O(m))^2$ which can be roughly as small as $\alpha(n)^2$. It is also worth pointing out that G is bipartite and all costs and capacities are either 1 or ∞ , so the hardness applies even in this restricted case.

4 A $2(n - 1)$ -approximation algorithm

It was first observed by Phillips [20] that there is always an optimal solution that is a subset of some cut $\delta(C)$. If we can identify a good cut to attack then only a Knapsack problem stands in our way. Accordingly, the crux of the Network Flow Interdiction problem is finding the right cut to attack. The strategy of the algorithm presented here is, at a basic level, to remove all of the edges whose capacity-to-cost ratio is very low, and then find a minimum cost s - t -cut in the resulting graph. The factor of 2 arises because we actually compare against a well structured 2-approximate solution rather than an optimal solution; it is an artifact of the “Knapsack portion” of the problem.

Let us begin illustrating the main idea with a simpler $(n - 1)$ -approximation algorithm for the special case that all interdiction costs are 1. We are given a graph $G = (V, E)$ with capacities $u : E \rightarrow \mathbb{N}$ and a budget $B \in \mathbb{N}$ as well as the vertices $s, t \in V$.

The first step is to order the edges by capacity so that $u(e_1) \leq u(e_2) \leq \dots \leq u(e_m)$ with ties broken arbitrarily. There is an optimal solution that chooses some cut $C^* \subseteq V$ and removes the B arcs with the highest labels, i.e., highest capacities, from the cut $\delta(C^*)$. Let R^* be this set of arcs and let $j^* = \max\{j \mid e_j \in \delta(C^*) \setminus R^*\}$ be the identity of a highest capacity edge in this cut that is not removed. We can “guess” j^* simply by trying each of the m possibilities. Let $E_{\leq} = \{e_j \in E \mid j \leq j^*\}$ and $E_{>} = E \setminus E_{\leq}$, and create two subgraphs of G , $G_{\leq} = (V, E_{\leq})$ and $G_{>} = (V, E_{>})$. The optimal interdiction solution is a budget feasible s - t cut in $G_{>}$, and it is a s - t cut of capacity OPT in G_{\leq} . The edges of these two cuts partition $\delta(C^*)$.

Now, another guessing step. We guess $f^* = \{w^*, v^*\}$ to be a pair of vertices separated by C^* that has the highest value minimum w - v -cut in G_{\leq} , among all those pairs of vertices w, v separated by C^* . Let C_{f^*} denote any minimum capacity w^* - v^* -cut in G_{\leq} . Since C^* is a w^* - v^* -cut we have by minimality of C_{f^*} that

$$\text{OPT} = u(\delta_G(C^*) \setminus R^*) = u(\delta_{\leq}(C^*)) \geq u(\delta_{\leq}(C_{f^*})),$$

which means that the capacity of C_{f^*} in G_{\leq} is a lower bound on the optimal objective value. Next, we contract in $G_{>}$ every pair of vertices $\{w, v\}$ that have minimum w - v -cut capacity in G_{\leq} larger than $u(\delta_{\leq}(C_{f^*}))$. Let $G'_{>}$ be the resulting graph and let C be a minimum cost s - t -cut in $G'_{>}$. Notice that, by the definition of f^* , no pair of vertices on opposing sides of the budget in feasible cut C^* gets contracted. Therefore, since C^* is a budget feasible cut in $G_{>}$, it follows that $R = \delta(C) \cap E(G'_{>}) = \delta(C) \cap E_{>}$ is budget feasible by its minimality. R is the set returned by the algorithm.

It remains to prove that R presents an $(n - 1)$ -approximation; for this, a Gomory-Hu tree is helpful. Recall that a Gomory-Hu tree for a graph $H = (V, E)$, with edge capacities u , is a tree T on vertices V along with weights κ on the edges of T . The edge weights have the property that, for any $w, v \in V$ and any minimum weight edge e on the unique w - v path in T , the two connected components of $T - e$ describe a minimum capacity w - v -cut in H , and the capacity of the cut is equal to $\kappa(e)$. A Gomory-Hu tree always exists and one can be found in $O(n^3 \sqrt{m})$ time (see, e.g., [17]).

Consider the cut C in the original graph G . The edges of $\delta(C) \setminus R$ are the same as those in $\delta_{\leq}(C)$. For every pair $w, v \in V$ that is not contracted, the capacity of a minimum capacity w - v -cut in G_{\leq} is bounded above by $u(\delta_{\leq}(C_{f^*})) \leq \text{OPT}$. In order to bound the capacity of $\delta_{\leq}(C)$ it suffices to show that it can be covered by a few minimum cuts. Obviously, $|\delta_{\leq}(C)|$ minimum cuts suffices, but this may be arbitrarily large. We can do better.

Lemma 5. *For any undirected, capacitated graph $H = (V, E)$ and any cut $C \subseteq V$, there is a set P of at most $|V| - 1$ pairs of vertices and, for each $\{w, v\} \in P$, a minimum capacity w - v -cut C_{wv} such that $\delta(C) \subseteq \cup_{p \in P} \delta(C_p)$.*

Proof. Let T be a Gomory-Hu tree for H and take $P = \delta_T(C)$. For each $\{w, v\} \in P$ the cut C_{wv} is the cut formed by deleting the edge wv from T and taking either of the two connected components. $|P| \leq |V| - 1$ because T is a tree, and each C_{wv} is a minimum w - v -cut because T is a Gomory-Hu tree. For each $w'v' \in \delta_H(C)$, there is an edge $wv \in \delta_T(C)$ on the unique w' - v' -path in T . Thus $w'v' \in C_{wv}$, which completes the proof. \square

Lemma 5 is easily seen to be sharp. If H is a star the cut that separates the center from the leaves cannot be covered by fewer than $n - 1$ minimum cuts.

Applying Lemma 5 to the output of the algorithm, we have

$$u(\delta(C) \setminus R) \leq \sum_{p \in P} u(\delta_{\leq}(C_p)) \leq (n - 1)u(\delta_{\leq}(C_{f^*})),$$

where the last inequality follows from the contraction step. On the other hand, f^* crosses C^* by definition, so $u(\delta_{\leq}(C_{f^*})) \leq \text{OPT}$ because it is the weight of a minimum cut separating its endpoints in G_{\leq} . Combining these inequalities we have $u(\delta(C) \setminus R) \leq (n - 1)\text{OPT}$, which completes the analysis of the algorithm in the special case that all of the costs are 1. The Gomory-Hu tree is also useful for implementing the algorithm because it is an efficient means to organize the computation of the minimum cut values and to organize the iteration over all of the possible pair-wise minimum cut values.

For general costs, the optimal solution can be interpreted as first choosing a cut C^* to attack and then interdicting an optimal set of edges in $\delta(C^*)$. Notice that the problem of interdicting an optimal set of edges in $\delta(C^*)$ is a Knapsack problem, which can also be interpreted as a Knapsack cover problem, when we want to select an optimal set of edges in $\delta(C^*)$ *not* to be interdicted. From the Knapsack Cover perspective, the set of edges in the cut that remains after interdiction is a minimum capacity subset of $\delta(C^*)$ with cost at least $c(\delta(C^*)) - B$. This formulation is convenient for us because we are interested in approximating the remaining capacity of the cut after interdiction.

In the $c = 1$ case, we were able to exploit the simple structure of the interdiction solution within $\delta(C^*)$ in order to guess j^* , the maximum capacity edge that is not removed by an optimal solution. We used the fact that the edges remaining in $\delta(C^*)$ after interdiction are those with lowest capacity in order to split G into G_{\leq} and $G_{>}$. Optimal Knapsack Cover solutions are much more complex, but we can exploit a much simpler structure when focusing on approximate Knapsack Cover solutions. The point is that it is nearly optimal to remove the edges with the highest ratio of capacity to cost, as long as the total cost of the edges removed is a sizable fraction of the budget.

Let us formalize that idea. In the Knapsack Cover problem, we are given as input a set of items $E = \{e_1, e_2, \dots, e_m\}$ with values $u : E \rightarrow \mathbb{N}$, costs $c : E \rightarrow \mathbb{N}$, and minimum expenditure B' . The goal is to output a minimum value subset $S \subseteq E$ such that $c(S) \geq B'$. Let

$$\rho(e) := u(e)/c(e)$$

be the *efficiency* of item e .

Fact 6. *The following algorithm is an $O(m^2)$ time 2-approximation algorithm for Knapsack Cover.*

Algorithm 1 A $2(n-1)$ -approximation algorithm for NFI.

Input: $G = (V, E)$, $s, t \in V$, $u : E \rightarrow \mathbb{N}$, $c : E \rightarrow \mathbb{N}$

Output: $R \subseteq E$ with $c(R) \leq B$ and $\max\text{-s-t-flow}(G - R) \leq (n-1)\text{OPT}$.

- 1 Order the edges so that $\rho(e_1) \leq \rho(e_2) \leq \dots \leq \rho(e_m)$.
 - 2 For each $j = 1, 2, \dots, m$ and $e \in E$
 - 3 Let $E_{\leq} = \{e_i \mid i \leq j \text{ and } u(e_i) \leq u(e)\}$ and $E_{>} = E \setminus E_{\leq}$.
 - 4 Construct a Gomory-Hu tree (T, κ) for (V, E_{\leq}) with capacities u .
 - 5 For each f in $E(T)$
 - 6 Form $G'_{>}$ by contracting in $(V, E_{>})$ all pairs of vertices w, v such that $wv \in E(T)$ with $\kappa(wv) > \kappa(f)$.
 - 7 Find a minimum cost s - t -cut C in $G'_{>}$.
 - 8 Store $R_{j,f,e} = \delta(C) \cap E_{>}$ if $c(R_{j,e,f}) \leq B$.
 - 9 Return the stored set that minimizes $\max\text{-s-t-flow}(G - R_{j,e,f})$.
-

- 1 *Guess f^* , the highest value element in some optimal solution and add it to S .*
- 2 *Remove from consideration all other items e with $u(e) \geq u(f^*)$.*
- 3 *Greedily add the remaining items with lowest efficiency to S until $c(S) \geq B'$.*
- 4 *Return S .*

It is well known that guessing the k highest value elements in an optimal Knapsack Cover solution, rather than just 1, leads to a $(1 + \frac{1}{k})$ -approximation algorithm at the expense of extending the running time to $O(m^{1+k})$. We make use of this improvement later.

The $2(n-1)$ -approximation algorithm for general costs operates in much the same way as the algorithm for the special case $c = 1$, described above. The main difference is that we will compare against the 2-approximation algorithm that finds an optimal cut C^* to attack and attacks it with the 2-approximation algorithm from Fact 6. Doing so involves one extra guessing step. The full procedure is Algorithm 1.

Theorem 7. *Algorithm 1 is a $2(n-1)$ -approximation algorithm for NFI. The running time is $O(m^{5/2}n^3)$.*

Proof. Consider the 2-approximation algorithm that optimally selects a cut C^* and determines a set of edges to remain after interdiction with the Knapsack 2-approximation algorithm in Fact 6. Let $R' \subseteq E$ be the interdiction set returned by this algorithm. We have $\text{OPT}' := \max\text{-s-t-flow}(G - R') \leq 2\text{OPT}$ by Fact 6. Let j^* be the index of the highest efficiency edge that remains in the cut $\delta(C^*)$, let $e^* \in \delta(C^*) \setminus R'$ be the highest capacity edge that remains in the cut $\delta(C^*)$, and let f^* be the highest weight edge of the Gomory-Hu tree T , constructed in the algorithm for $j = j^*$ and $e = e^*$, that crosses C^* . It is sufficient to show that $c(R_{j^*,e^*,f^*}) \leq B$ and

$$\max\text{-s-t-flow}(G - R_{j^*,f^*,e^*}) \leq (n-1)\text{OPT}',$$

since the solution returned by the algorithm will be budget feasible and have objective value no greater than this one.

Consider the iteration when $j = j^*$, $e = e^*$, and $f = f^*$. By the definitions of j^* , e^* , and R' we have $R' = \delta(C^*) \cap E_{>}$. Since R' is budget feasible, this means C^* is a budget feasible

s - t -cut in $G_{>} = (V, E_{>})$ and is an OPT' -capacity s - t -cut in $G_{\leq} = (V, E_{\leq})$. In step 6, no pair of vertices on opposing sides of C^* are contracted because f^* is a highest weight edge of T crossing C^* . Because of this, $\delta(C^*) \cap E_{>}$ remains as the edge set of a (budget feasible) cut in $G'_{>}$ which implies that R_{j^*, e^*, f^*} , as the edge set of a minimum cost s - t -cut, is budget feasible.

By Lemma 5 we have $u(\delta(C) \cap E_{\leq}) \leq (n-1)\kappa(f^*)$. Finally, $\kappa(f^*)$ is the capacity of a minimum cut in G_{\leq} that separates the endpoints of f^* . So our observation that G_{\leq} contains a s - t -cut C^* of capacity OPT' that separates the endpoints of f^* implies $\kappa(f^*) \leq \text{OPT}'$. Combining the inequalities we have, as desired, $\max\text{-}s\text{-}t\text{-flow}(G - R_{j^*, e^*, f^*}) \leq 2(n-1)\text{OPT}$.

The algorithm requires constructing no more than m^2 Gomory-Hu trees and m^2 minimum s - t -cut computations, so the running time is dominated by the Gomory-Hu trees, which take $O(\sqrt{mn}^3)$ time each. Thus the total time taken by Algorithm 1 is $O(m^{5/2}n^3)$. \square

As we mentioned earlier, the Knapsack Cover 2-approximation algorithm of Fact 6 can be converted into a $(1 + \frac{1}{k})$ -approximation algorithm by guessing k edges rather than just one. If we add the extra guessing to Algorithm 1 then we can improve the factor 2 in its approximation ratio, arriving at the following corollary to Theorem 7.

Corollary 8. *For any integer $k > 0$, there exists a $(1 + \frac{1}{k})(n-1)$ -approximation algorithm for NFI that runs in time $O(m^{\frac{3}{2}+k}n^3)$.*

5 Variants of NFI

A natural question related to NFI is to ask whether there is an approximation algorithm that approximates the change in the maximum flow, rather than the residual flow in the network. We call NFI with this alternative objective the reduction NFI problem. An α -approximation for NFI is generally not an approximation algorithm for reduction NFI, since an α -approximation is not enough to determine whether the maximum reduction in flow is zero or positive. So, we cannot count on Algorithm 1 to approximate reduced NFI to within any factor, and it turns out that achieving any multiplicative approximation ratio is NP-hard.

It was shown in [24] that it is NP-hard to determine whether the optimal objective value is 0 or positive for the reduction version of the Bipartite Matching Interdiction problem, where edges are removed to reduce the cardinality of a maximum matching. We can easily reduce Bipartite Matching Interdiction to NFI with the standard reduction from bipartite maximum matchings to flows. Thus we have the following.

Proposition 9. *It is NP hard to determine whether the optimal objective for reduction NFI is 0 or positive.*

The main consequence of Proposition 9 is that there is no multiplicative approximation algorithm for reduction NFI.

The Budgeted Minimum s - t -Cut problem, BMstC, is also essentially the same as NFI. In the following α may depend on n but we drop the dependence to simplify the notation.

Theorem 10. *If there exists a time $T(n, m)$ α -approximation algorithm for NFI then there is a time $O(T(n, 2m))$ α -approximation for BMstC. If there exists a time $T(n, m)$ α -approximation algorithm for BMstC then, for any integer $1 \leq k$, there exists a time $O(m^{1+k}T(n, m))$ α' -approximation for NFI, where $\alpha' = (1 + \frac{1}{k})\alpha$.*

Proof. Let $G = (V, E), u, c, s, t, B$ be an instance of BMstC. Let E_1 and E_2 be disjoint copies of E and construct a NFI instance on the graph $G' = (V, E_1 \cup E_2)$ as follows. Given the two copies $e_1 \in E_1$ and $e_2 \in E_2$ of $e \in E$ assign NFI costs c' and capacities u' as follows: $c'(e_1) = c(e)$, $u'(e_1) = \infty$, $c'(e_2) = \infty$, and $u'(e_2) = u(e)$. Now, for any s - t -cut $C \subseteq V$, $c(\delta_G(C)) \leq B$ if and only if $R = \delta_{G'}(C) \cap E_1$ satisfies $c'(R) \leq B$ and the minimum s - t -cut in $G - R$ has capacity $u'(\delta_{G'}(C) \setminus R) = u'(\delta_{G'}(C) \cap E_2) = u(\delta_G(C))$. This establishes an objective preserving 1-1 correspondence where each budget feasible cut in G is matched with a budget feasible NFI solution that has as its edge removal set all finite cost edges of G' crossing the cut. The first claim follows immediately from this 1-1 correspondence.

Now suppose $G = (V, E), u, c, s, t, B$ is an instance of NFI and order the edges $E = \{e_1, \dots, e_m\}$ according to increasing efficiency $\rho(e) = u(e)/c(e)$. To solve this instance with a BMstC algorithm, we first guess k edges that remain in the attacked cut after an optimal interdiction and one that is interdicted. Let R^* be an optimal interdiction set that attacks a s - t -cut $C^* \subseteq V$. We will guess the set $S^* \subseteq \delta(C^*) \setminus R^*$ of $\min\{k, |\delta(C^*) \setminus R^*|\}$ edges crossing C^* with the highest capacities that remain after interdiction with R^* and the edge $f^* \in R^*$ with least label (i.e. least efficiency) of any edge in R^* . There are no more than $m \sum_{i=0}^k \binom{m-1}{i} \leq m^{k+1}$ guesses S, f to try.

For each guess S, f and each edge $e \in E$, we adjust the cost and capacity of e as follows. If $e \in S$ or if it has a label lower than that of f , set the cost of e to 0 and leave its capacity unchanged. Otherwise, that is if $e \notin S$ has its label at least as large as the label of f , set its capacity to 0 and leave its cost unchanged. When $S = S^*$ and $f = f^*$ the optimal BMstC solution corresponds to a $(1 + \frac{1}{k})$ -approximate NFI solution, which follows by an extension of Fact 6 upon comparing against the appropriate Knapsack Cover problem on $\delta(C^*)$. Since cost $b \leq B$ cuts with capacity ν for the BMstC instance correspond exactly to cost b interdiction solutions with objective value at most ν , an α -approximate budgeted minimum cut corresponds to a α' -approximate NFI solution.

Solving NFI in this manner requires running the BMstC-algorithm once for each of the $O(m^{1+k})$ guesses and additional $O(n + m)$ overhead for creating the auxiliary graphs, capacities, and costs. Thus the total time is $O(m^{1+k}T(n, m))$. \square

The proof of Theorem 10 shows that NFI manifests itself as the combination of a BMstC and a Knapsack problem. Notice that the Knapsack problem is indeed a very special case of NFI. For this consider an NFI problem on a graph that has only two vertices s and t with many edges between them. There is only one s - t -cut, so to solve BMstC one only needs reduce the value of this unique s - t -cut as much as possible through interdiction. This is clearly a knapsack problem.

Another variant we want to mention briefly is *min-cost NFI*. In this variant we impose that $\max\text{-}s\text{-}t\text{-flow}(G - R) \leq B$ and the goal is to minimize $c(R)$. In fact, NFI and min-cost NFI are equivalent, as has been shown in [25], and this carries over to approximation factors.

Finally, we mention two other variants of NFI. One allows for directed edges and the other allows for interdiction by vertex removals, instead of or in addition to edge removals. We call these directed NFI and node-wise NFI, respectively. Clearly, each of these is at least as hard as NFI, by replacing each edge with anti-parallel arcs in the directed case and by subdividing each edge with a vertex in the node-wise case. It is also possible to modify the reduction used in Theorem 3 slightly to adapt it to each of these settings. Now, combining the material in this section we have the following corollary to Theorem 3.

Corollary 11. *If there is an $\alpha(n)$ -approximation algorithm for the directed, node-wise, or min-cost variants of NFI or BMstC, then there is a $2\alpha(n^2)^2$ -approximation algorithm for DkS .*

In Corollary 11, we have absorbed the extra $(1 + \epsilon)$ that appears in the reduction from NFI to BMstC into the 2. This is possible with the assumption that $k \geq 3$ upon revisiting the bound in (1).

Finally, we remark that Joret and Vetta [14] have proved that reducing the rank of a transversal matroid is DkS -hard to approximate. Recall that a transversal matroid is defined by a bipartite graph $(X \cup Y, E)$, its ground set is X and a set $S \subseteq X$ is independent if the graph has a matching that covers S . The rank reduction problem is to find a minimum cardinality set $R \subseteq X$ such that the maximum matching with no endpoints in R is below a given threshold. Their results do not imply hardness for NFI (or more appropriately, min-cost NFI), but there are hardness implications for some of the variants. By taking the standard formulation of Maximum Bipartite Matching as a Network Flow problem, their Theorem 3.4 implies that any $O(n^\epsilon)$ -approximation algorithm for node-wise min-cost NFI or for directed min-cost NFI can be used for an $O(n^{4\epsilon})$ -approximation to the densest k -subgraph. However, this is a weaker statement than Corollary 11 and their approximation ratio does not improve if DkS has hard sparse instances, whereas ours does as we discussed at the end of Section 3.

References

- [1] N. Assimakopoulos. A network interdiction model for hospital infection control. *Computers in Biology and Medicine*, 17(6):413–422, 1987.
- [2] C. Bazgan, S. Toubaline, and D. Vanderpooten. Complexity of determining the most vital elements for the p-median and p-center location problems. *Journal of Combinatorial Optimization*, 25(2):191–207, 2013.
- [3] C. Bazgan, S. Toubaline, and D. Vanderpooten. Critical edges for the assignment problem: Complexity and exact resolution. *Operations Research Letters*, 41(6):685–689, 2013.
- [4] C. Bazgan, S. Toubaline, and D. Vanderpooten. Critical edges/nodes for the minimum spanning tree problem: complexity and approximation. *Journal of Combinatorial Optimization*, 26(1):178–189, 2013.
- [5] A. Bhaskara, M. Charikar, E. Chlamtac, U. Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: an $O(n^{1/4})$ approximation for densest k -subgraph. In *Proceedings of the forty-second ACM Symposium on Theory of Computing*, pages 201–210, 2010.
- [6] A. Bhaskara, M. Charikar, A. Vijayaraghavan, V. Guruswami, and Y. Zhou. Polynomial integrality gaps for strong SDP relaxations of Densest k -Subgraph. In *Proceedings of the twenty-third annual ACM-SIAM Symposium on Discrete Algorithms*, pages 388–405, 2012.
- [7] C. Burch, R. Carr, S. Krumke, M. Marathe, C. Phillips, and E. Sundberg. A decomposition-based pseudoapproximation algorithm for network flow inhibition. In *Network Interdiction and Stochastic Integer Programming*, pages 51–68. Springer, 2003.

- [8] Julia Chuzhoy, Yury Makarychev, Aravindan Vijayaraghavan, and Yuan Zhou. Approximation algorithms and hardness of the k -route cut problem. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 780–799, 2012.
- [9] M. Dinitz and A. Gupta. Packing interdiction and partial covering problems. In *Integer Programming and Combinatorial Optimization*, pages 157–168. Springer, 2013.
- [10] U. Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the thirty-fourth annual ACM Symposium on Theory of Computing*, pages 534–543, 2002.
- [11] U. Feige, D. Peleg, and G. Kortsarz. The dense k -subgraph problem. *Algorithmica*, 29(3):410–421, 2001.
- [12] Greg N Frederickson and Roberto Solis-Oba. Increasing the weight of minimum spanning trees. *Journal of Algorithms*, 33(2):244–266, 1999.
- [13] T.E. Harris and F.S. Ross. Fundamentals of a method for evaluating rail net capacities. Technical Report RM-1573, RAND Corp. 1955.
- [14] G. Joret and A. Vetta. Reducing the rank of a matroid. *arXiv:1211.4853*, 2012.
- [15] Leonid Khachiyan, Endre Boros, Konrad Borys, Khaled Elbassioni, Vladimir Gurvich, Gabor Rudolf, and Jihui Zhao. On short paths interdiction problems: total and node-wise limited interdiction. *Theory of Computing Systems*, 43(2):204–233, 2008.
- [16] S. Khot. Ruling out PTAS for graph min-bisection, dense k -subgraph, and bipartite clique. *SIAM Journal on Computing*, 36(4):1025–1071, 2006.
- [17] B. Korte and J. Vygen. *Combinatorial Optimization*. Springer, 2002.
- [18] A. T. Murray, T. C. Matisziw, and T. H. Grubescic. Critical network infrastructure analysis: interdiction and system flow. *Journal of Geographical Systems*, 9(2):103–117, 2007.
- [19] C. H Papadimitriou and M. Yannakakis. On the approximability of trade-offs and optimal access of web sources. In *Proceedings of the forty-first annual Symposium on Foundations of Computer Science*, pages 86–92, 2000.
- [20] C. A. Phillips. The network inhibition problem. In *Proceedings of the twenty-fifth annual ACM Symposium on Theory of Computing*, pages 776–785, 1993.
- [21] J. O. Royset and R. K. Wood. Solving the bi-objective maximum-flow network-interdiction problem. *INFORMS Journal on Computing*, 19(2):175–184, 2007.
- [22] A. Schrijver. On the history of the transportation and maximum flow problems. *Mathematical Programming*, 91(3):437–445, 2002.
- [23] R. K. Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, 17(2):1–18, 1993.
- [24] R. Zenklusen. Matching interdiction. *Discrete Applied Mathematics*, 158(15):1676–1690, 2010.

- [25] R. Zenklusen. Network flow interdiction on planar graphs. *Discrete Applied Mathematics*, 158(13):1441–1455, 2010.
- [26] R. Zenklusen. An $O(1)$ -approximation for minimum spanning tree interdiction. In *Proceedings of 56th Annual IEEE Symposium on Foundations of Computer Science*, 2015. To appear.